

Intracranial Study of Speech-Elicited Activity on the Human Posterolateral Superior Temporal Gyrus

Mitchell Steinschneider^{1,2}, Kirill V. Nourski³, Hiroto Kawasaki³, Hiroyuki Oya³, John F. Brugge^{3,4,5} and Matthew A. Howard III³

¹Department of Neurology, ²Department of Neuroscience, Albert Einstein College of Medicine, New York 10461, NY, USA, ³Department of Neurosurgery, University of Iowa, Iowa City, IA 52242, USA, ⁴Department of Psychology and ⁵Department of Physiology, University of Wisconsin, Madison, WI 53706, USA

Address correspondence to Dr Mitchell Steinschneider, Department of Neurology, Rose F. Kennedy Center, Room 322, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, NY, USA. E-mail: mitchell.steinschneider@einstein.yu.edu.

To clarify speech-elicited response patterns within auditory-responsive cortex of the posterolateral superior temporal (PLST) gyrus, time–frequency analyses of event-related band power in the high gamma frequency range (75–175 Hz) were performed on the electrocorticograms recorded from high-density subdural grid electrodes in 8 patients undergoing evaluation for medically intractable epilepsy. Stimuli were 6 stop consonant–vowel (CV) syllables that varied in their consonant place of articulation (POA) and voice onset time (VOT). Initial augmentation was maximal over several centimeters of PLST, lasted about 400 ms, and was often followed by suppression and a local outward expansion of activation. Maximal gamma power overlapped either the N α or P β deflections of the average evoked potential (AEP). Correlations were observed between the relative magnitudes of gamma band responses elicited by unvoiced stop CV syllables (/pa/, /ka/, /ta/) and their corresponding voiced stop CV syllables (/ba/, /ga/, /da/), as well as by the VOT of the stimuli. VOT was also represented in the temporal patterns of the AEP. These findings, obtained in the passive awake state, indicate that PLST discriminates acoustic features associated with POA and VOT and serve as a benchmark upon which task-related speech activity can be compared.

Keywords: average evoked potentials, electrocorticogram, gamma oscillations, place of articulation, voice onset time

Introduction

Multiple methodologies are required to clarify the neural mechanisms associated with speech processing. Studies of a possible hierarchy of cytoarchitectonic fields involved in speech perception have been generally modeled after the organization of auditory fields observed in nonhuman primates (Kaas and Hackett 2000, 2005; Hackett 2003, 2007, 2008). This organization envisions a hierarchical scheme of interconnected core, belt, and parabelt fields. Lateral belt fields exhibit tonotopic organizations that parallel those of adjacent core fields and a caudal belt area. Belt fields may be topographically organized with respect to band-passed noise (Morel et al. 1993; Rauschecker and Tian 2000, 2004) and the rate of frequency-modulated (FM) sweeps (Tian and Rauschecker, 2004), and several exhibit selectivity to the direction of a sound source in space (Rauschecker and Tian 2000; Tian et al. 2001; Woods et al. 2006). Belt areas may be highly responsive to conspecific vocalizations (Rauschecker and Tian 2000; Tian et al. 2001), though there remains controversy as to whether these areas are selectively responsive to specific call types (Recanzone 2008). While this model is appealing, it has been difficult on

anatomical grounds alone to define, outside of core areas, homologies between auditory cortical fields in the monkey and those in the human (Galaburda and Sanides 1980; Rivier and Clarke 1997; Hackett et al. 2001; Hackett 2003, 2007, 2008; Wallace et al. 2002; Sweet et al. 2005; Fullerton and Pandya 2007; Hackett 2007).

Previous human intracranial studies from our laboratory and others have identified an acoustically responsive area on posterolateral superior temporal gyrus (STG) that we refer to as the posterolateral superior temporal auditory field, or PLST (Howard et al. 2000). Field PLST responds to both simple and complex sounds, including speech, with robust averaged evoked potentials (AEPs), the latter further being modulated by audiovisual interactions (Celesia 1976, Crone et al. 2001; Brugge et al. 2003; Reale et al. 2007; Besle et al. 2008; Bidet-Caulet et al. 2008; Chang et al. 2010).

Functional neuroimaging studies have shown activation of a region of the STG that overlaps with electrophysiologically identified PLST. As found in electrophysiological studies of PLST, functional magnetic resonance imaging (fMRI) results show activity arising here being evoked by language-related stimuli, as well as by a wide range of speech and nonspeech sounds (Binder et al. 2000; Hall et al. 2002; Scott and Johnsrude 2003; Hart et al. 2004; Uppenkamp et al. 2006). As a rule, however, more complex sounds, such as FM or amplitude-modulated (AM) tones, syllables, or words produce greater activation of this area than do simpler sounds such as noise bursts or pure tones. Results of imaging studies have led investigators to conclude that this region on the STG represents an intermediate stage in the neural networks subserving speech perception (Scott and Johnsrude 2003; Liebenthal et al. 2005; Uppenkamp et al. 2006; Poeppel et al. 2008). If this turns out to be the case, then it is not at all clear whether this area should be considered an auditory belt or parabelt field in the current hierarchical model or whether it should be considered an area that has arisen later in human evolution (Hackett 2003, 2007).

Regardless of any homology PLST might share with non-human primates, its functional organization remains poorly understood. Based on maps of AEPs obtained by click stimulation, it was suggested earlier that PLST may not be functionally homogeneous (Howard et al. 2000). Neuroimaging studies have not yet identified functional organizational patterns within this area, possibly because of limits of spatial resolution imposed by this methodology.

This paper addresses questions related to the role of PLST in processing fundamental features of speech sounds, using simple consonant–vowel (CV) syllables presented in a passive

listening paradigm. In these experiments, we have employed time-frequency analysis of event-related band power (ERBP) in the electrocorticogram (ECoG) recorded directly from the cortex, which allows identification of “induced” changes in the ECoG that, in contrast to the AEP, are not phase-locked to the stimulus. Using this approach, it has been shown that sound evokes prominent changes in the high gamma frequency range (>75 Hz) of the ECoG recorded in both Heschl’s gyrus (HG, Brugge et al. 2009; Nourski et al. 2009) and PLST (Crone et al. 2001, 2006; Edwards et al. 2005, 2009, 2010; Towle et al. 2008) that provide information not observed in the AEP. In general, activity in the gamma frequency band has been shown to reflect important features in sensory perception (Herrmann et al. 2004; Fries et al. 2007; Jensen et al. 2007). Furthermore, augmented power in the gamma frequency range correlates both with fMRI results and with neuronal firing patterns (Mukamel et al. 2005; Niessing et al. 2005; Ray, Crone, et al. 2008; Privman et al. 2007). Finally, gamma augmentation recorded near lamina 1 has been shown to be a sensitive and reliable index of the A1 tonotopic organization as defined by multiunit activity recorded in lower lamina 3 in the awake monkey (Steinschneider et al. 2008). This latter finding suggests that organizational patterns of gamma activation recorded above the pial surface in the human brain may also be a meaningful reflection of unit activity organization located in the subjacent cortical depths. Taken together, the above considerations indicate that time-frequency analysis of the ECoG may provide valuable insights into the organization of PLST as it relates to speech perception.

The goals of this paper are to 1) define the temporal and spatial dynamics of syllable-evoked activity at spatial resolutions higher than those used previously in similar studies (e.g., Crone et al. 2001; Edwards et al. 2009, 2010) in order to better understand basic characteristics of PLST activation and 2) identify whether ERBP patterns exist on PLST that promote the neural representation of 2 of the most fundamental phonetic parameters of speech; stop consonant place of articulation (POA) and voice onset time (VOT). Results of this study will thus help clarify the role PLST plays in the neural network of language using ERBP (see Canolty et al. 2007; Edwards et al. 2010).

Materials and Methods

Subjects

Eight subjects (4 women) aged 20–56 years participated in this study. Each patient was being evaluated for potential surgical amelioration of medically intractable epilepsy. All research protocols were approved by the University of Iowa Human Subjects Review Board prior to their initiation, and all patients gave informed consent prior to their participation. Subjects could rescind consent at any time during the research procedures without interrupting their medical evaluation. Research recordings did not disrupt the simultaneous acquisition of clinically required data, and the patients did not incur any additional risks by their study participation. Subjects remained on a varying regimen of anticonvulsant drugs at the discretion of their neurologist. Each subject had normal hearing as defined by standard audiometric studies, except subject L145, who had a moderate sensorineural hearing deficit at frequencies >3 kHz. However, all subjects, including L145, exhibited normal speech discrimination scores. Additional subject demographics are presented in Table 1. Clinical evaluations indicated that neither HG nor PLST were seizure foci in any patient.

Table 1

Subject summary

Hemisphere of recording and subject ID	Age and sex	Handedness	Wada (language dominance)	Verbal IQ	Full scale IQ
R139	53 F	Left	Right	76	77
R142	33 F	Right	Left	80	82
L145	56 F	Right	Left	91	95
L146	29 F	Right	Left	89	98
L147	29 M	Left	Left	105	100
R149	22 M	Left	Bilateral	82	75
R175	20 M	Right	Left	Not available	111
L178	47 M	Right	Left	103	96

Stimuli

Six stop CV syllables were presented (/ba/, /da/, /ga/, /pa/, /ta/, and /ka/). These syllables had been used previously, thus facilitating comparisons of evoked neural activity in auditory-responsive areas on the lateral temporal lobe with that in HG (Steinschneider et al. 1999) and monkey primary auditory cortex (A1) (Steinschneider et al. 1995, Steinschneider and Fishman 2010). Syllables were constructed on the parallel branch of a KLSYN88a speech synthesizer (Klatt DH and Klatt LC 1990), contained 4 formants (F1 through F4), and were 175 ms in duration (Fig. 1). Fundamental frequency began at 120 Hz and fell linearly to 80 Hz. Steady-state formant frequencies were 700, 1200, 2500, and 3600 Hz. Onset frequencies for F2 and F3 of /ba/ were 800 and 2000 Hz, 1600 and 3000 Hz for /da/, and 1600 and 2000 Hz for /ga/. Formant transitions for F2 and F3 were 40 ms in duration. All syllables had the same F1, which had an onset frequency of 200 Hz and a 30 ms formant transition, and F4, which did not contain a formant transition. Thus, /ga/ had the same F2 as /da/ and the same F3 as /ba/. A 5-ms period of frication exciting F2–F4 preceded the onset of voicing. Amplitude of frication was increased by 18 dB at the start of F2 for /ba/ (800 Hz) and /ga/ (1600 Hz) and F3 for /da/ (3000 Hz). Spectral structure was such that /ba/ and /da/ had diffuse onset spectra maximal at either low or high frequencies, whereas /ga/ had a more compact onset spectrum maximal at intermediate values. These patterns have been proposed to be important for discrimination of stop consonants varying in their POA (e.g., /b/ vs. /d/ vs. /g/) (Stevens and Blumstein 1978). The unvoiced CV syllables (/pa/, /ta/, and /ka/) were identical to their voiced counterparts (/ba/, /da/, and /ga/) except for an increase in the VOT from 5 to 40 ms. For the unvoiced CV syllables, the next 35 ms contained aspiration noise. Syllables were accurately identified during informal presentations to multiple members of the laboratories at Albert Einstein College of Medicine and University of Iowa and are available upon request.

Recording Methods

Clinically targeted arrays of platinum–iridium disc electrode contacts embedded in a silicon membrane were implanted on the pial surface overlying the temporal lobe (Howard et al. 2000; Reale et al. 2007). Electrode contacts were 1.6 mm in diameter and had an inter-contact spacing of 5 mm. The grid arrays were centered over more posterior temporal regions and coverage typically included the superior, middle, and portions of the inferior temporal gyri, with additional extension that variably included segments of parietal, central, and frontal brain areas. Grid arrays contained 96 (12 × 8) electrodes in 7 subjects and 64 (8 × 8) in another. Additional subdural arrays and depth electrodes were also implanted as clinical needs dictated. Depth electrodes were also implanted into the ipsilateral HG of the 4 subjects with right hemisphere arrays. These hybrid depth electrodes (Howard et al. 1996) had 4 platinum circumferential macro-contacts, 0.5 mm in length and spaced 10 mm apart, with more closely spaced micro-contacts situated between them. Recordings from the depth electrodes are beyond the scope of this paper.

Pre- and postimplantation MRI scans were performed in all subjects (Siemens TIM Trio 3 T with 12 channel head array coil). Structural MRI volumes were created from T_1 -weighted images (coronal scans, 256 × 256 matrix with 1-mm slice thickness, 0.78 × 0.78 × 1.0 mm in-plane

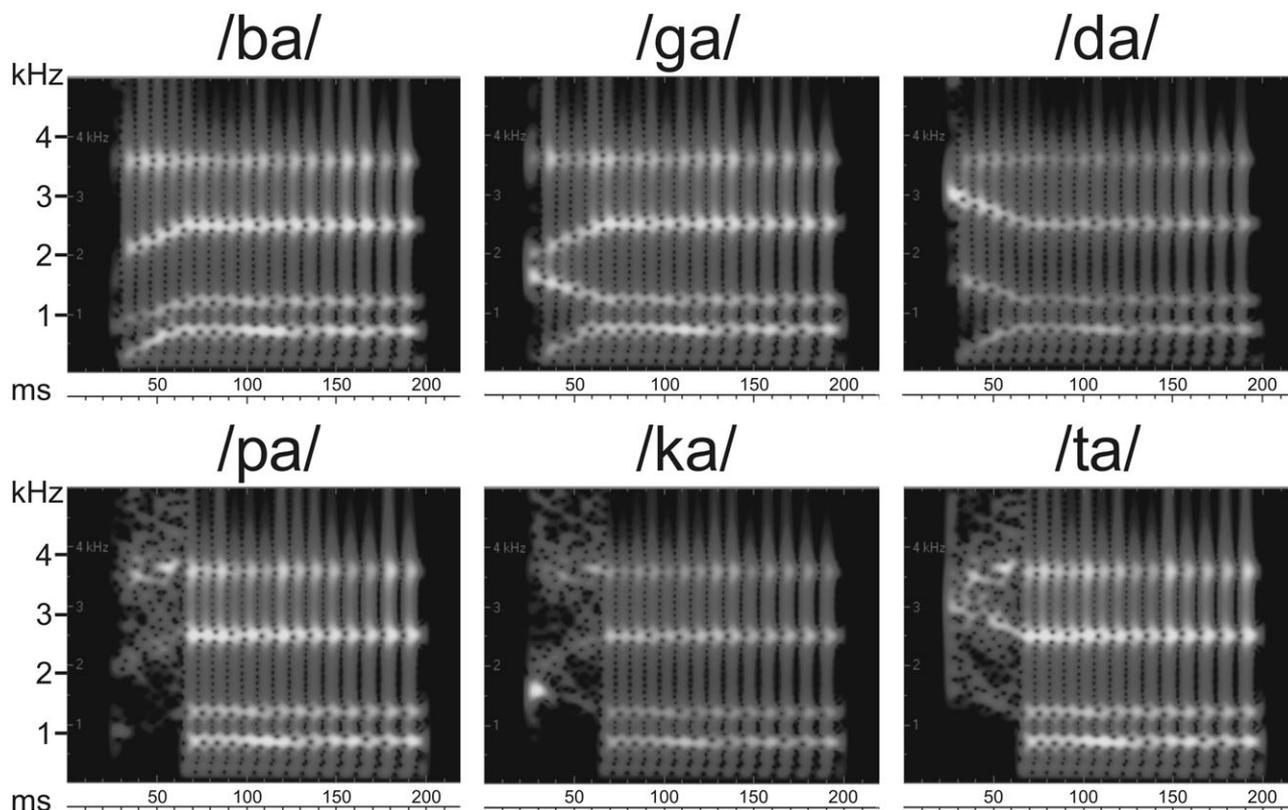


Figure 1. Synthetic syllables used in the present study. Amplitudes of the higher formants are emphasized in the figure. On average, the steady-state portions of F1 were 8.5 dB louder than F2, with further fall-offs of 3.4 and 4.8 dB between higher formants. See text for additional details.

resolution with 2 image set averaging) using Analyze software (version 9, Analyze Direct). High resolution intraoperative photographs were used to identify grid contact locations relative to gross anatomical landmarks on the hemispheric surface. Electrode contact positions were then manually depicted on matching cortical locations identified on the preoperative surface-rendered MR image. Depth electrode contact locations were identified on the postimplantation MRI and depicted on the preoperative image set using a manual anatomical template matching technique (Brugge et al. 2008, 2009).

Experimental recording sessions began several days after surgical implantation. Data acquisition was performed in electrophysiological recording suites in the University of Iowa General Clinical Research Center. Subjects were awake and either sitting comfortably in their hospital bed or in a nearby chair. However, their attentional states were not strictly controlled for during this recording paradigm. Thus, sounds were presented with the subjects awake but not actively attending to the stimuli, and intermittent covert attention cannot be excluded. The ECoG was acquired simultaneously from all electrodes at a sampling rate of 2034.5 Hz and with a bandpass of 1.6–1000 Hz using a TDT RX5 or RZ2 processor (Tucker-Davis Technologies). The recording reference was a platinum disc electrode contact from an additional grid array located on the undersurface of the ipsilateral anterior temporal lobe or in contact with the galea near the vertex of the skull. These contacts were relatively “inactive” relative to the arrays overlying the lateral temporal lobe.

Syllables were delivered binaurally through calibrated insert earphones (ER4B; Etymotic Research), which were integrated into custom-fitted ear molds. Sound intensity levels were individually chosen by each subject to provide for comfortable listening and easy discrimination of the syllables. These intensities were ~50 dB above hearing threshold. All 6 syllables were presented with equal probability in random order in the same recording blocks with an interstimulus-onset interval of 2 s in 6 subjects. In subjects L178 and R175, the voiced (/ba/, /ga/, and /da/) and unvoiced (/pa/, /ka/, /ta/) CV syllables were

presented in sequential recording blocks. Typically, each stimulus was presented 50 times.

Data Analysis

ECoG data were analyzed in the time domain by calculating the AEPs and in the time-frequency domain by computing the ERBP. ERBP analysis of individual stimulus trials in the recording blocks was performed within the range of 10–250 Hz in 5 Hz increments using a wavelet transform based on complex Morlet wavelets (Oya et al. 2002). The wavelet constant ratio was defined as $2\pi f_0 \sigma = 9$, where f_0 is the center frequency of the wavelet and σ is its standard deviation in frequency. ERBP was calculated on a trial-by-trial basis as a ratio relative to median baseline power within the same 5 Hz frequency band measured between 100 and 200 ms prior to stimulus onset. Edge effects (contribution of energy from artifacts induced by the windowing of the recording epochs to the poststimulus onset interval used to estimate baseline power) were negligible for the range of center frequencies that corresponded to high gamma frequency bands. Individual trials with waveform peaks/troughs greater than 2.5 standard deviations from the mean of the 50 stimulus presentations were rejected from analysis to minimize contamination by electrical artifacts, epileptic spikes or paroxysmal high amplitude slow-wave activity. The lead author examined the results of the reject calculations to further minimize the possibility that abnormal activity contaminated the recordings and to ensure that this reject value did not lead to the discarding of high amplitude signals time-locked to stimulus presentations.

Average ERBP was calculated in 200 ms nonoverlapping bins beginning at stimulus onset and extending for 1 s, and 50 ms nonoverlapping bins beginning at stimulus onset and extending for 300 ms. Both time scales are thought to be important for the analysis of speech (e.g., Boemio et al. 2005; Poeppel et al. 2008). The rationale for using 200 ms bins was also based on 1) examining gamma activity in a time frame that would provide a broad overview of syllable-evoked

activity and be more compatible with the temporal resolution of fMRI, 2) examining gamma activity in a time frame such that the first bin would overlap the 3 principal AEP deflections ($P\alpha$, $N\alpha$, and $P\beta$) generated on PLST (Howard et al. 2000), and 3) examining gamma activity such that the first bin would overlap nearly the entire duration of the stimuli and not include activity evoked by the offset of the sounds (based on a response latency on PLST > 25 ms). The rationale for examining gamma activity at the higher temporal resolution of 50 ms bins was also based on 1) preliminary analysis showing that no significant activity occurred within the first 50 ms after stimulus onset, and thus the first time bin effectively isolated this latency of response period, and 2) this window of integration is relevant for differentiating activity based on consonant POA as defined by multiunit activity in the monkey and AEPs in the human (Obleser et al. 2006; Tavabi et al. 2007; Steinschneider and Fishman 2010; see also Chang et al. 2010 who used a 40 ms window).

Average ERBP was first computed for the 5 Hz wide frequency bands. Subsequently, changes were averaged together to produce 3 bands: 75–90, 95–115, and 125–175 Hz. The purpose of binning together the 5 Hz frequency bands into larger bands (e.g., 75–90 Hz) was to minimize the overcontribution that the highest frequencies (e.g., 125–175 Hz) would have if the 5 Hz bands were simply averaged together. The power in each of these 3 bands at each electrode contact in the grid arrays was computed in each of the time bins and averaged together for each stimulus and for each subject. The resultant values were then analyzed to determine if they were significantly different from baseline using t -tests corrected for multiple comparisons by the false discovery method of Benjamini and Hochberg (1995) (see also Boatman-Reich et al. 2010). Following this analysis, the responses to all syllables were averaged together to compute an average syllable-evoked response at each electrode contact. All nonsignificant changes were treated as zeros in the averaging across all syllables. Lower gamma frequencies (30–70 Hz) were excluded to minimize contributions from faster frequency components in the AEP. Furthermore, a comparison analysis that included 30–45 and 50–70 Hz bands was carried out in the same manner to determine if the inclusion of only high gamma activity resulted in more robust response changes. A paired t -test comparing the 30–175 Hz range (5 bands) versus the 75–175 Hz range (3 bands) revealed that the more restrictive 75–175 Hz range elicited larger responses when examining ERBP within the first 200 ms time bin ($t_{727} = 10.86$, $P < 0.0001$). Additional methodological details will be presented in the Results where appropriate to minimize redundancy.

Results

Syllables Elicit Large Amplitude Responses from Locations on PLST

Large amplitude AEPs and ERBP were recorded from multiple electrode contacts located on PLST of both hemispheres and in all subjects (for one data set, see Supplementary Fig. 1). These responses are exemplified by the AEPs and ERBP from an electrode contact located on PLST in subject L146 in response to the stimulus /pa/ (Fig. 2A). The AEP shown in Figure 2B is characterized by a polyphasic waveform with an initial positive deflection having a peak around 50 ms ($P\alpha$), a large negative deflection around 160 ms ($N\alpha$) followed by a large positive wave with peak latency around 235 ms (see Howard et al. 2000). ERBP derived from the same ECoG activity represented in the AEP is shown in Figure 2C. Pronounced power increases spanning a wide frequency range are elicited. Maximal power is centered in the high gamma bands and extends upward to the highest frequencies analyzed.

Distribution of Initial Activation Overlying Temporal Cortex

Significant gamma activation within the first 200 ms after stimulus onset extends along much of PLST. Data from each of

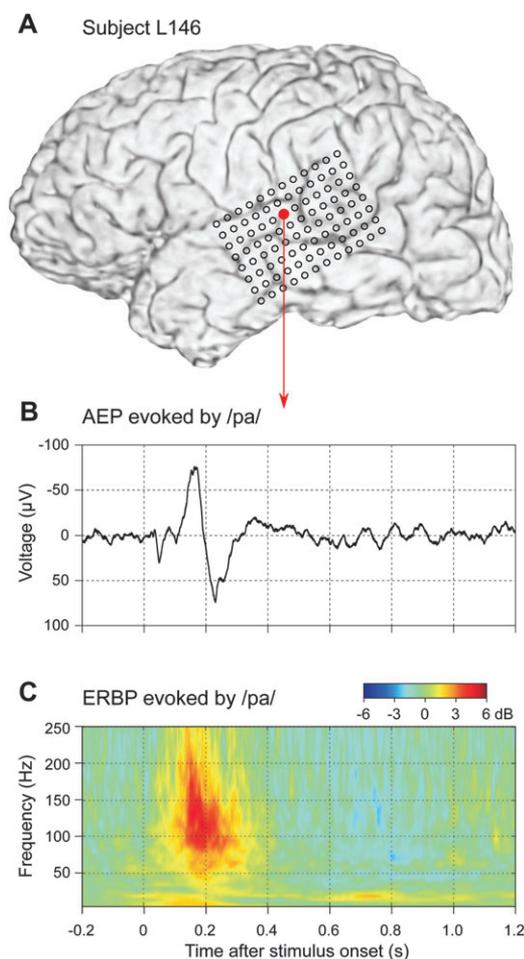


Figure 2. Example of AEP and ERBP recorded from a site on PLST in subject L146 in response to the syllable /pa/. (A) Electrode grid superimposed on image of the left hemisphere. The electrode contact from which the exemplar data were recorded is shown in red. (B) AEP evoked by /pa/. The initial positive wave peaks at 50 ms and is followed by major negative and positive deflections that peak at 167 and 235 ms, respectively. (C) ERBP elicited by /pa/. Maximal power increases are centered in the high gamma bands.

the 8 subjects are shown in Figures 3 and 4 for left and right hemisphere grids, respectively. Color coding is based on the ratio of mean ERBP for all 6 syllables within the first 200 ms after stimulus onset compared with baseline, expressed in dB. Filled circles represent recording sites where increases in gamma activity were significantly greater than baseline for at least one syllable. Color represents the magnitude of the power change. Larger filled circles indicate that significant differences occurred for all 6 syllables; smaller filled circles indicate that significant differences occurred for fewer than 6. Open circles depict sites at which gamma activity failed to reach significance threshold. Dysfunctional electrode contacts are designated by an “x” on the grid. Sites exhibiting larger increases in power, and those exhibiting significant increases to all 6 syllables, are largely restricted to posterior STG, while lower amplitude increases and sites where not all syllables elicited significant increases are often observed over adjacent cortical regions.

Temporal Dynamics of Gamma Activity

Many electrode sites exhibiting high-amplitude gamma activity during the first 200 ms following syllable onset

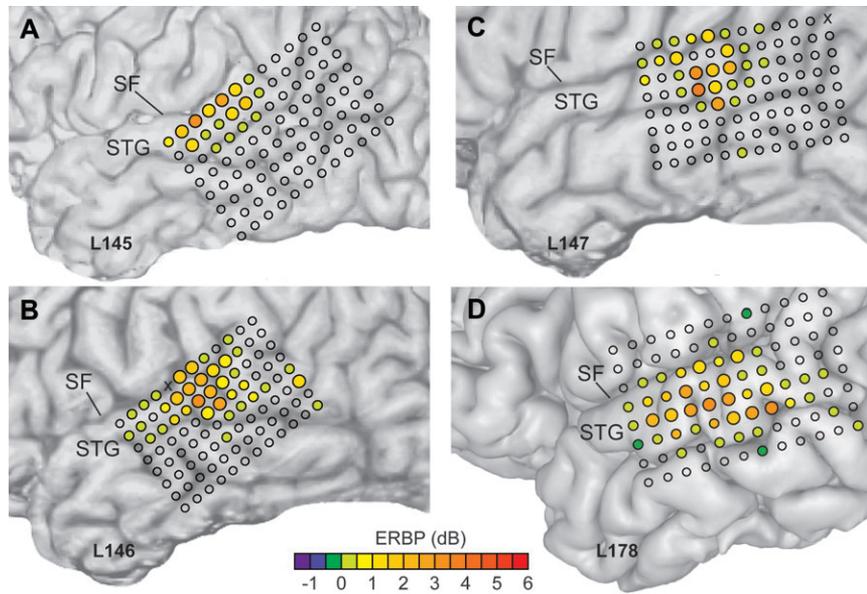


Figure 3. ERBP at high gamma frequencies elicited by the averaged response to the 6 syllables in the 4 subjects with left hemisphere electrode grids. Responses are measured from 0 to 200 ms after stimulus onset and represent the ratio between the ERBP during this time period and baseline. Larger colored circles represent electrode sites where all 6 syllables elicited responses significantly greater than baseline. Smaller colored circles represent sites where at least one syllable elicited a significant response. Dysfunctional electrodes are marked with an “x.” The largest increases in ERBP are restricted to circumscribed regions of PLST. Electrode spacing is 5 mm. SF denotes the Sylvian fissure.

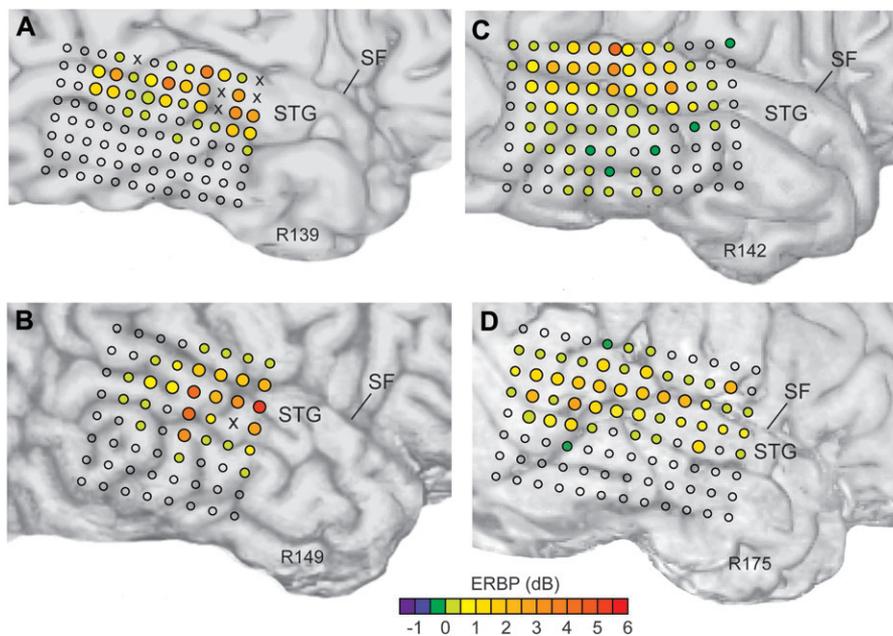


Figure 4. ERBP at high gamma frequencies elicited by the averaged response to the 6 syllables in the 4 subjects with right hemisphere electrode grids. Conventions for the illustration are the same as for Figure 3.

continue to do so during the 200- to 400-ms time interval. This pattern might have been further extended for sounds longer than the 175 ms syllables, but this speculation will require future investigation using sounds of varying durations. After the initial period of activation, most sites displaying maximal gamma activity show a pronounced decrement of activity that often falls below prestimulus baseline levels, usually beginning in the 400–600 ms time interval. While the specific time values of increases and decreases in ERBP varied across subjects, this basic pattern

was observed for all subjects. The change from augmentation to suppression of the gamma activity is illustrated in Figure 5, which depicts the overall grid patterns of high gamma power averaged across all 6 syllables from 0 to 600 ms for subjects L145 (A–C) and R142 (D–F). For both subjects, power increases that are maximal in the first 200 ms diminish over time. At many sites, power in the 400- to 600-ms time interval falls significantly below baseline levels. This effect is most pronounced at the 2 sites in each subject that had maximal excitation during the first 200 ms (arrows).

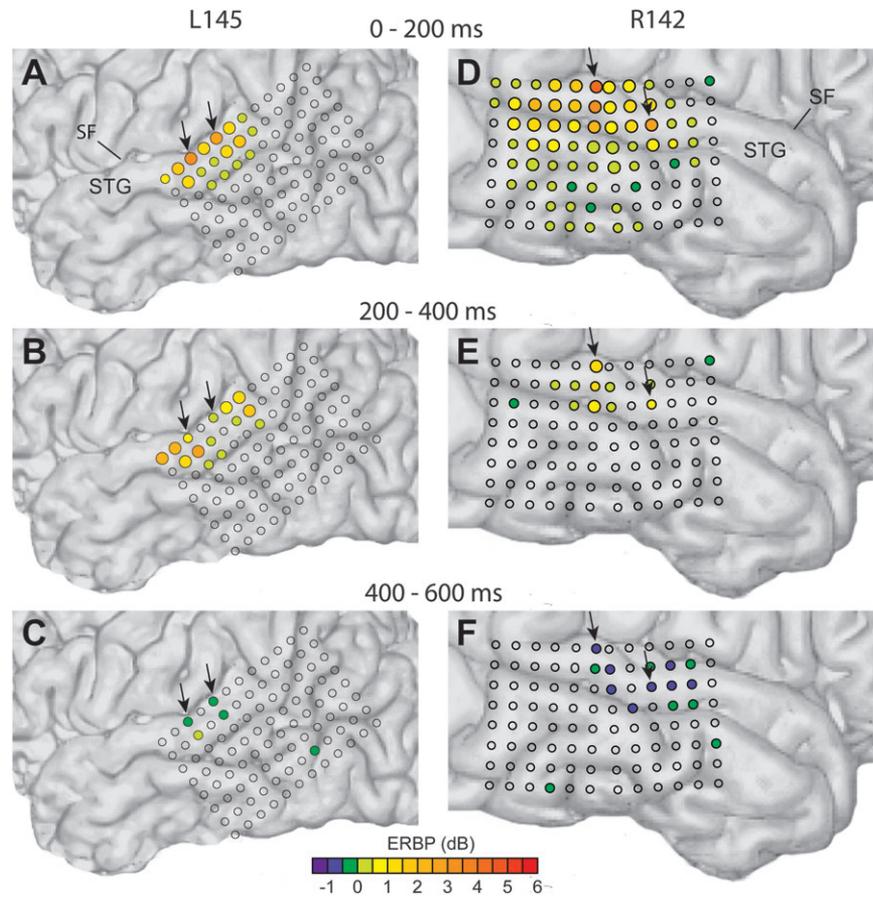


Figure 5. Temporal patterns of gamma activity from 0 to 600 ms for subjects L145 (A–C) and R142 (D–F) are shown in the left and right columns, respectively. Electrode contacts showing maximal activity in the 0- to 200-ms time interval show activity significantly below baseline levels by the 400- to 600-ms time interval (arrows).

Examined at 50-ms intervals, averaged ERBP is generally maximal between 100 and 200 ms. Maximal power changes occur between 100 and 150 ms in 2 subjects, 150 and 200 ms in 5 subjects, and 200 and 250 ms in 1 subject. When compared against the temporal profile of the AEP, these ERBP maxima overlap either the latter half of the large initial negative deflection ($N\alpha$) or the leading half of the following positivity ($P\beta$). These patterns are illustrated in Figure 6.

Figure 6 depicts the AEP averaged across all syllables from -100 to 500 ms poststimulus onset for each subject. The AEPs chosen for illustration represent the waveform for each subject exhibiting the largest initial negative deflection ($N\alpha$). The AEP for subject R142 also displays short-latency deflections evoked by the offset of the stimuli. Plotted with the AEPs are the mean ERBPs across the entire grids in 50-ms intervals from 0 to 300 ms (filled circles). These time intervals are marked by the vertical dotted lines in the figure. Almost without exception, there are no significant power changes evoked by the syllables within the first 50 ms after stimulus onset. Maximal ERBP always overlaps with either the falling phase of $N\alpha$ or the rising phase of $P\beta$. There is also a trend for the peak latency of the ERBP to parallel the latency of the AEP. Thus, the 2 subjects (R142, R149) with the shortest latency peaks in the ERBP also have among the shortest $N\alpha$ deflections of all the subjects. In contrast, the subject (L178) with the longest latency peak in the ERBP has the longest latency $N\alpha$ deflection for all the subjects. Intermediate relationships are observed for the other 5 subjects.

Spatial Dynamics of Gamma Activity

The time course of gamma activation along PLST is not uniform, as there is an outward expansion of gamma augmentation from areas of initial activation to areas both anterior and posterior along the gyrus. Commonly, regions of maximal activation beginning soon after stimulus onset are replaced later in time by maximal activation at adjacent areas more anterior and posterior along PLST. This spatial extension of activity is illustrated in Figures 7 and 8. Figure 7 depicts the time course of activity in 50-ms intervals for 2 subjects. For L145, initial activation begins in central regions of PLST (A) and over time spreads in both anterior and posterior directions along the gyrus (B–C). Locations showing strong initial activation include those marked by arrows. Over time, there is a sequential spatial redistribution of ERBP such that by the 250- to 300-ms interval, those locations displaying maximal initial power increments show markedly decayed power, while surrounding areas contain pronounced increases in power. For R175, strong initial activation includes a site located over more posterior portions of PLST (D). Over several hundred milliseconds, this activation decays to a point where no syllable elicits significant increases in ERBP. Instead, areas anterior and posterior along PLST become the major sources of regional activation. Additional patterns of spatial dynamics of activity within PLST are shown in Supplementary Figure 2.

These dynamic changes are further illustrated at finer temporal resolution for selected electrode channels (Fig. 8).

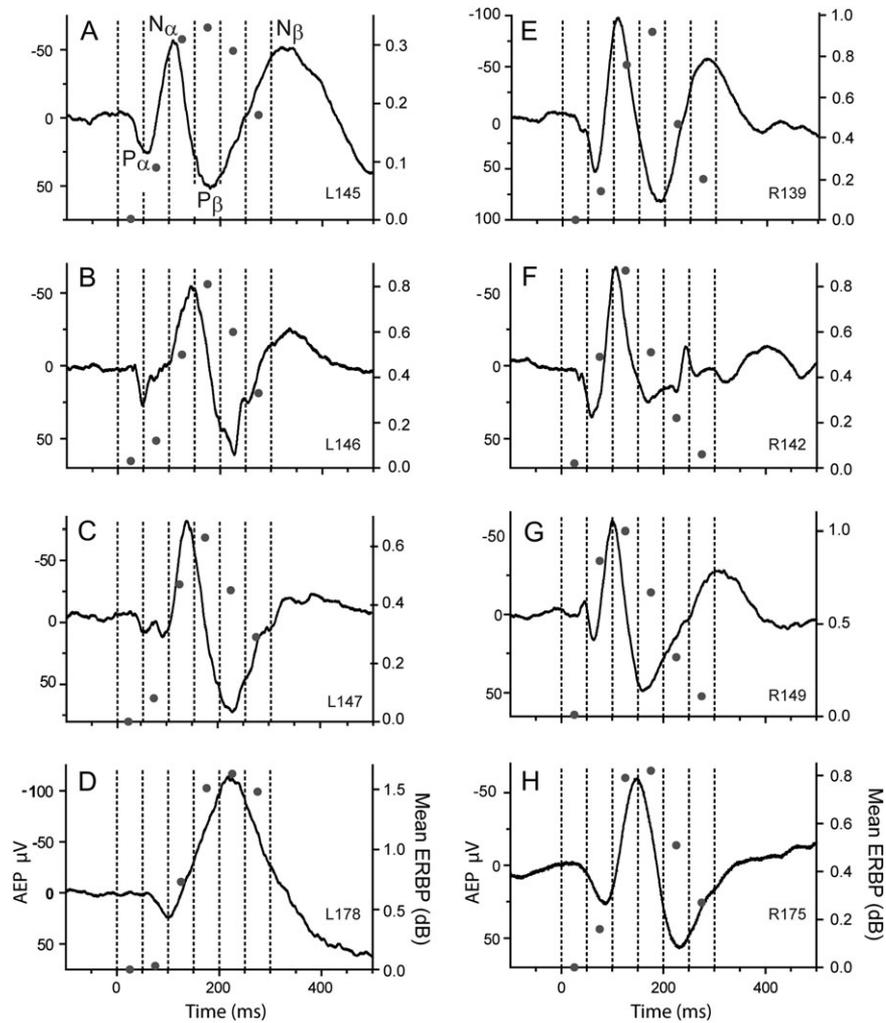


Figure 6. Temporal dynamics of ERBP plotted at 50-ms intervals (0–300 ms) superimposed on the AEPs with the largest initial negative component ($N\beta$) recorded from each subject. AEP amplitudes are denoted on the left y-axis, while the right y-axis depicts the ERBP averaged across all grid electrode sites and all 6 syllables. Vertical dotted lines demarcate the six 50-ms intervals. Major AEP components are labeled in the upper left-hand graph according to previously used conventions (Howard et al., 2000). See text for details.

The left-hand column (A–C) depicts electrode grids overlying PLST of 3 subjects, with selected sites identified by color-coded letters. To the right are ERBP waveforms derived from activity at these sites averaged across all 6 syllables. For L145 (A), activation is initially maximal at sites c and d, while those electrodes extending in 5 mm increments anterior (b and a) and posterior (e and f) along the grid from c and d have progressively longer onset latencies and more pronounced activity at later time periods.

Similar patterns are seen for L178 (B). Here, sites d (anterior) and f (posterior) have almost identical onset latencies and time courses and each represent the 2 sites with the largest early activation in the 50- to 100-ms time interval. Sites along PLST more anterior to d have later onsets, even though activity at site b is larger than that at site d. Later responses are observed at electrodes e (more anterior) and g (more posterior) relative to f, even though once again site e with a later onset has greater ERBP increases. Comparable spatial patterns are observed for subject R175 (C).

Representation of Consonant POA

To this point, focus has been on general patterns of gamma activity averaged across responses to all 6 CV syllables. This

analysis, however, does not examine whether speech-evoked activity is systematically modulated by the phonetic parameters of the consonants. To further explore this issue, we analyzed whether the relative magnitude of ERBP varied systematically as a function of the POA of the CV syllables.

Analyses were performed separately for the time intervals of 50–100, 100–150, and 150–200 ms. The interval of 0–50 ms was excluded based on minimal high gamma responses on PLST during this time (Fig. 6). Later intervals were excluded as it was reasoned that POA differences should occur early in the responses and later activity would likely reflect the common vowel shared by all syllables (Steinschneider and Fishman 2010). Electrode sites were included in the analyses if they contained significant responses to at least 4 of the speech sounds in the specified time intervals. This criterion minimized inclusion of false positive data (e.g., only one stimulus elicited a significant response), limited but did not exclude inclusion of sites not located on PLST (Supplementary Fig. 3), and ensured that at least one significant response from both the voiced (/b/, /g/, /d/) and unvoiced (/p/, /k/, /t/) consonants was included. The analysis began by ranking at each brain site studied the magnitudes of ERBP elicited by the 3 voiced CV syllables (/ba/,

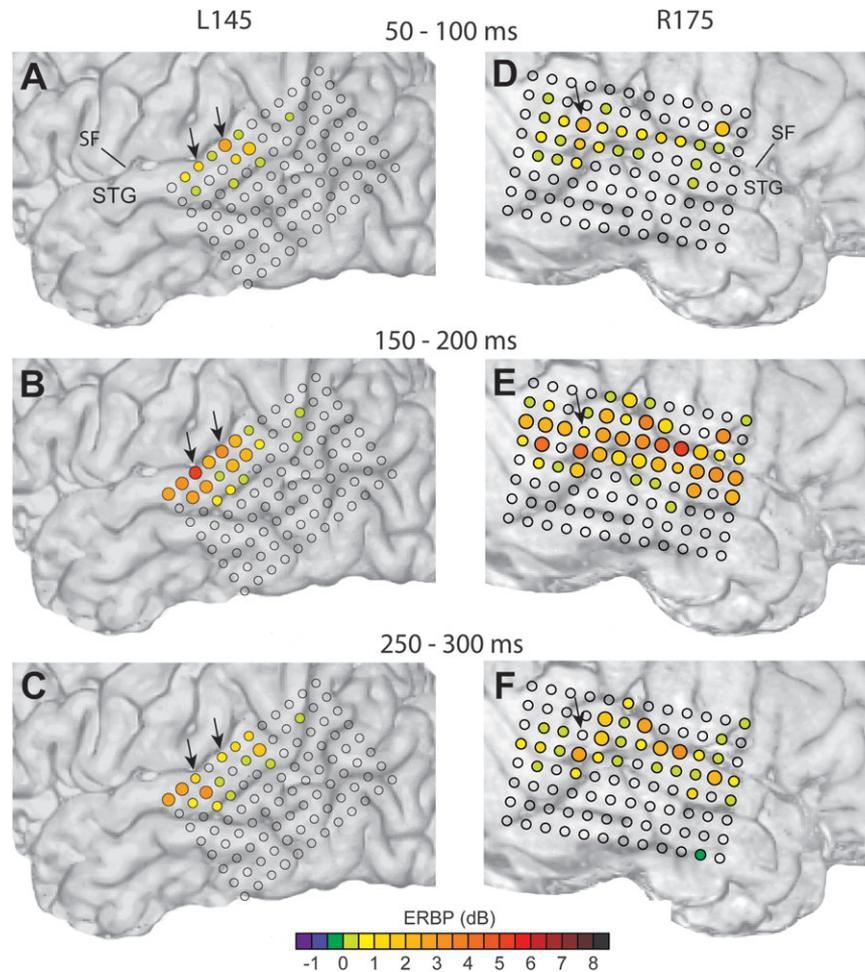


Figure 7. Spatial dynamics of ERBP plotted at 50-ms interval in the 3 time periods noted in the figure for subjects L145 (left column) and R175 (right column). For L145, there is a spatial movement of maximal activity from the center of PLST to regions more anterior and posterior along the gyrus (arrows). For R175, the electrode site showing maximal activity during the 50- to 100-ms time interval no longer shows significant activity at 250–300 ms (arrow), contrasting with surrounding sites more anterior and posterior along PLST.

/ga/, and /da/). The resulting rank order is a reflection of the selectivity of a given brain site to acoustic features associated with POA; the 3 stimuli having a common VOT. The created rank order pattern was then used to predict the relative amplitudes of the gamma responses elicited by their matched unvoiced CV syllables (/pa/, /ka/, and /ta/). If only one voiced CV syllable elicited a significant response, then the predicted ranking at that site was performed using the 3 unvoiced CV syllables and then compared with the responses elicited by the voiced CV syllables. It was reasoned that if the magnitudes of gamma band responses varied as a function of POA, independent of VOT, then the POA-based rank ordered responses should be comparable for the voiced and unvoiced stimulus sets. Cumulative statistical analyses (Friedman test) of the resulting rank orders were performed across all recording sites that reached the above-described inclusion criterion.

A significant correlation was found between the POA-based rankings of the responses to predicted and observed CV syllables that was restricted to the 100- to 150-ms time interval ($N = 174$, degrees of freedom = 2, Friedman statistic = 16.64, $P = 0.0002$). Results are shown in Figure 9. Post hoc tests (Dunn's multiple comparison test) revealed that rank 3 was greater than

both ranks 2 and 1 ($P < 0.01$). In contrast, the same test performed for the 50–100 ms ($N = 72$, Friedman statistic = 0.27) and 150–200 ms ($N = 209$, Friedman statistic = 2.42) time intervals did not yield significant differences. For the 100- to 150-ms time interval, correct predictions for the largest responses occurred at 44% of the brain sites examined (76/174, random ratio 33%, one-sided Fisher's Exact test $P = 0.03$). The most accurate predictions occurred when /ba/ or /pa/ gave the largest response (28/52, 54%). Predictions were correct when /da/ or /ta/ elicited the largest response at 42% of the brain sites examined (30/72), while the least accurate predictions occurred when /ga/ or /ka/ elicited the largest response (18/50, 36%). The latter result is likely based on the acoustic ambiguity of the synthetically constructed /ga/ and /ka/, which shared common second formant transitions with /da/ and /ta/ and third formant transitions with /ba/ and /pa/.

Significant correlations were also observed when comparing absolute response differences between relevant syllable pairs in the 100- to 150-ms time interval. Thus, comparisons between differences in responses evoked by /ba/ and /da/ (/ba/ - /da/) correlated with those simultaneously obtained between /pa/ and /ta/ (/pa/ - /ta/; Pearson correlation coefficient $r = 0.1795$, $P < 0.02$). These syllables have the most spectrally distinct POA

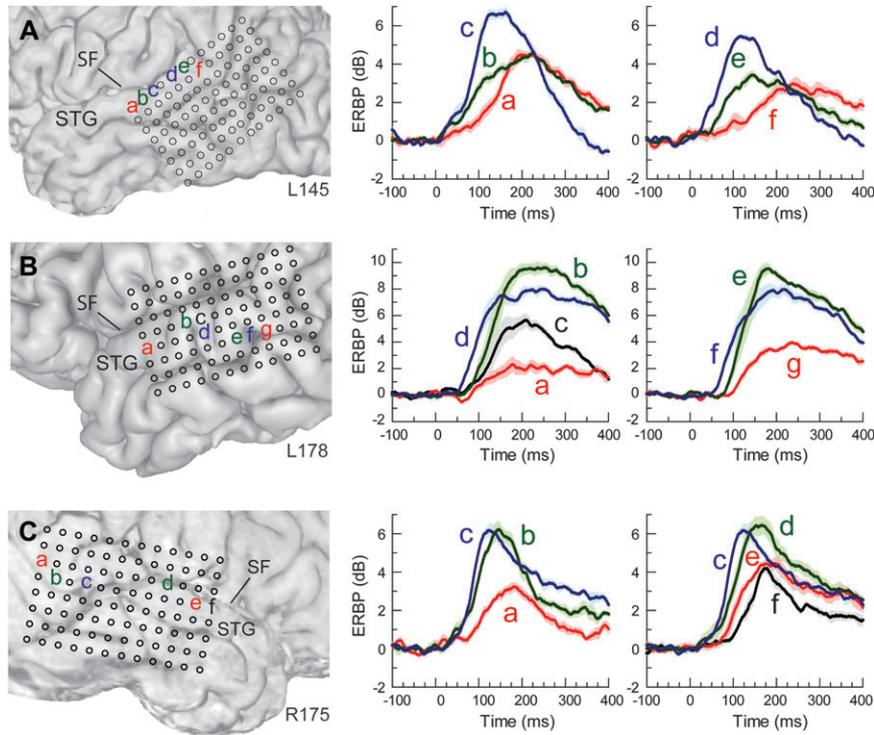


Figure 8. Spatial dynamics of ERBP plotted at high temporal resolution for subjects L145, L178, and R175. Selected electrode sites are color-coded superimposed on the 3 electrode grids. Power waveforms for these electrodes are shown in the center and right-hand columns. Waveforms represent the mean (\pm standard error) for the ERBP elicited by all 6 syllables. See text for details.

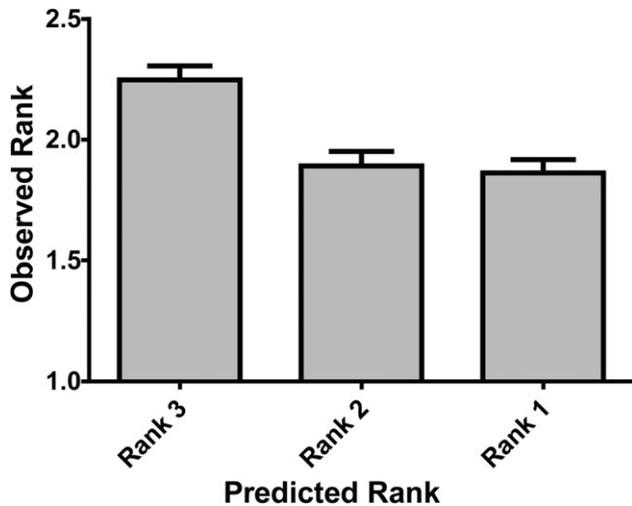


Figure 9. Rank order of gamma activity between 100 and 150 ms poststimulus onset for the voiced (VOT = 5 ms; /ba/, /ga/, and /da/) CV syllables partially predicts the rank order of gamma activity in the same time interval for the unvoiced CVs that have the same POA (VOT = 40 ms; /pa/, /ka/, and /ta/). Rankings are at the 174 sites that contained at least 4 significant responses to the syllables, thus guaranteeing that at least one significant response would occur for the voiced and unvoiced syllables. A significant correlation between the POA-based rankings of the responses to the voiced and unvoiced CV syllables was observed, thus indicating that PLST has an organization that respects, to some degree, the POA of stop consonants. See text for details.

profiles used in this study, with /ba/ and /pa/ dominated by lower frequencies at stimulus onset while /da/ and /ta/ are dominated by higher frequencies. Additionally, differences between responses evoked by /ga/ and /da/ correlated with

those simultaneously obtained between /ka/ and /ta/ (Pearson correlation coefficient $r = 0.1572$, $P < 0.04$). The correlation between responses elicited by /ba/ and /ga/ with those obtained for /pa/ and /ka/ was not significant.

Collectively, these findings demonstrate that the relative magnitude of gamma band responses within PLST are modulated as a function of acoustic stimulus features associated with stop consonant POA in the time interval of 100–150 ms after stimulus onset. However, because these analyses were performed on pooled data, the results cannot be used to identify possible topographic patterns of POA representation that may exist within the area of PLST from which recording were obtained. Further work is clearly required to identify if topographic patterns exist within PLST, to compare responses between dominant and nondominant hemispheres, and to relate speech-evoked activity with those obtained from simpler stimuli (e.g., pure, AM, and FM tones).

Representation of VOT

Previous studies have shown that VOT is partially represented by temporal patterns of activity within posteromedial HG, a region that corresponds to core auditory cortex (Steinschneider et al. 1999, 2005; Liégeois-Chauvel et al. 1999). Syllables with short VOTs evoke AEPs with a characteristic triphasic morphology, while syllables with a long VOT contain additional components time-locked to voicing onset. Furthermore, this change in response morphology based upon syllable VOT is more prominently observed in more anterolateral portions of the posteromedial HG (Steinschneider et al. 1999, 2005).

Similar patterns in the AEP recorded from HG were observed in the present study, but a full description of these data is beyond the scope of this paper. Shown, however, are

representative AEPs obtained from PLST in 2 subjects (R139, R149) averaged over responses to voiced (blue) and unvoiced (red) syllables (Fig. 10). AEP morphology is modulated by the VOT of the syllables. There are prominent shifts in the latencies of all but the initial peaks in the waveforms (elicited by stimulus onset) evoked by the unvoiced CV syllables. These shifts are always to later time points and approximate the 40 ms VOT of these syllables. New waveform deflections are introduced that approximate the extended VOT value. These VOT-modulated profile types are observed to varying extents in all 6 subjects, though there was a tendency for recordings from the right hemisphere to demonstrate the effect to a greater degree than from the left. Additional subjects will be required to further characterize this observation.

In addition to a temporal representation of VOT in PLST, there are also reliable differences in the magnitude of ERBP elicited by voiced versus unvoiced CVs. These findings are shown in Figure 11, which depicts the difference in amplitude of gamma activity from 100 to 150 ms for /ba/ and /pa/ (same POA, different VOT) compared with differences for /da/ and /ta/ (filled black circles, solid black regression line) and /ga/ and /ka/ (open gray squares, dashed gray regression line) at all grid electrode sites that had significant responses to at least 4 of the 6 syllables. These sites are the same as those used for analysis of POA previously described. For both VOT-based comparisons, there is a statistically significant relationship between the gamma responses elicited by the voiced versus unvoiced CV syllables. Pearson correlation coefficients comparing (/ba/ - /pa/) with (/da/ - /ta/) and (/ga/ - /ka/) are both 0.79 ($P < 0.0001$). Linear regression lines are also significantly different from zero for the 2 comparisons ($F_{1,172} = 294.3$, $P < 0.0001$ and $F_{1,172} = 287.2$, $P < 0.0001$, respectively). Correlations identified in the 100- to 150-ms time interval continued into the 150- to 200-ms time interval. Now, Pearson correlation coefficients comparing (/ba/ - /pa/) with (/da/ - /ta/) and (/ga/ - /ka/) are 0.68 and 0.69, respectively ($P < 0.0001$). In

contrast, correlations in the 50- to 100-ms time interval were not statistically significant.

The absence of a correlation for VOT in the 50- to 100-ms interval does not mean that this stimulus parameter did not modulate speech-elicited gamma activity during this time period, as many electrode sites failed to reach the criterion for inclusion into the analysis (i.e., significant responses from fewer than 4 CVs). When comparing the summed amplitudes of gamma activity across all grid electrode sites for each subject between the 3 voiced and the 3 unvoiced syllables and normalizing the responses to a percentage of the total power across all syllables in this time period, there was a 2-fold increase in power for the voiced versus the unvoiced syllables (0.67 vs. 0.33, paired t -test $t_7 = 7.301$, $P = 0.0002$). This ratio decreased during the 100- to 150-ms interval to 0.56 for the voiced syllables and 0.44 for the unvoiced syllables (paired t -test $t_7 = 1.312$, $P = 0.23$) and equalized during the 150- to 200-ms interval (0.50 for each).

Discussion

Summary of Findings

This study identifies 4 characteristics of changes in high gamma frequency power elicited by CV syllables as recorded from subdural grid electrodes placed over auditory area PLST. 1) Initial activation in the gamma frequency range is maximal over several centimeters of field PLST. Power is generally maximal between 100 and 200 ms following stimulus onset and overlaps the latter half of the $N\alpha$ and leading half of the $P\beta$ deflections of the AEP. Gamma activation lasts ~ 400 ms and is often followed by suppression below baseline levels. 2) There is an outward expansion in both anterior and posterior directions within PLST from a region of initial activation adjacent to HG on the lateral STG. 3) Correlations exist between the relative magnitudes of gamma band responses elicited by unvoiced

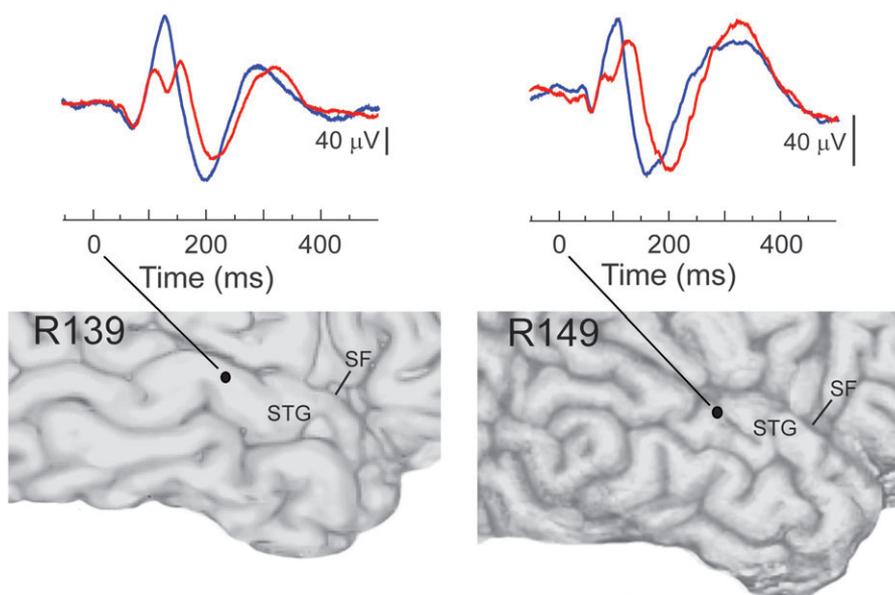


Figure 10. AEPs recorded from PLST are modulated by syllable VOT. AEPs evoked by the averaged responses to the voiced CV syllables (/ba/, /da/, and /ga/) are shown in blue, while the superimposed AEPs evoked by the averaged responses to the unvoiced CV syllables (/pa/, /ta/, and /ka/) are shown in red. Anatomical locations where AEPs were recorded for subject R139 and R149 are shown. Calibration bars shown to the right of each AEP denote 40 μ V of deflection (negative up-going). See text for details.

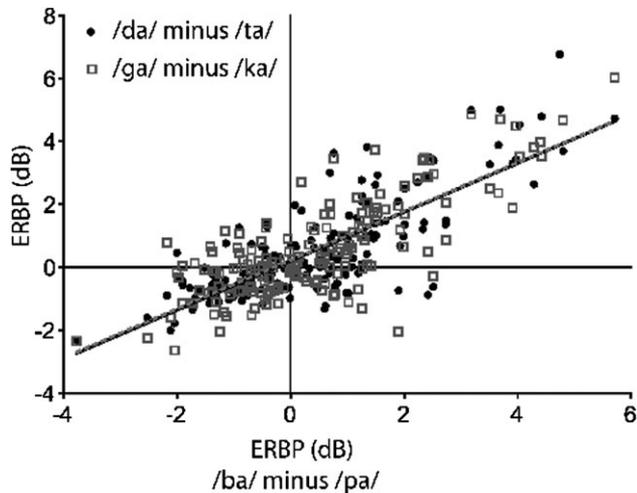


Figure 11. Reliable differences in the magnitude of gamma band responses elicited by voiced versus unvoiced CVs are observed on PLST. The difference in amplitude of gamma activity from 100 to 150 ms for /ba/ and /pa/ (same POA, different VOT) are plotted against the differences for /da/ and /ta/ (filled black circles, solid regression line) and /ga/ and /ka/ (open gray squares, dashed regression line) at all 174 grid electrode sites that showed at least 4 significant responses to the syllables (same as for Fig. 9). For both VOT-based comparisons, there is a statistically significant relationship between the gamma responses elicited by the voiced versus unvoiced CV syllables. See text for details.

stop CV syllables and their corresponding voiced stop CV syllables. 4) VOT has both a temporal and an amplitude representation in PLST, as observed in differences in AEP waveforms and relative magnitudes of gamma band responses. These latter findings indicate that PLST responses are modulated by acoustic speech features associated with the phonetic features of POA and VOT.

Distribution and Timing of Initial Activation Overlying Lateral Cortex

Numerous studies have demonstrated that speech sounds elicit short-latency high-frequency gamma responses from the posterior region of the STG (e.g., Crone et al. 2001, 2006; Canolty et al. 2007; Edwards et al. 2009, 2010). The spatial resolution of the grid electrodes used in the present study allows a more detailed estimation of the extent of this activation. While partially dependent on grid placement, maximal excitation, as defined by significant responses to all 6 syllables throughout most of the sounds' duration, spans from 2.5 to 4.5 cm along PLST, and lesser excitation can occur across the entire width of the grid (5.5 cm). Maximal excitation tends to hug the contours of PLST, indicating the relative specificity of the response. Because markedly different patterns can occur between adjacent sites (e.g., in amplitude, time course, and stimulus sensitivity), a spatial resolution for high gamma ERBP of approximately 0.5 cm (spacing of electrodes) can be estimated.

Maximal activation occurs between 100 and 200 ms following stimulus onset, with responses significantly larger than baseline beginning at about 50 ms. Concordant values have been reported previously (Crone et al. 2001; Edwards et al. 2005, 2009, 2010; Canolty et al. 2007). While the extended duration of gamma augmentation (~400 ms) has also been reported (Crone et al. 2001; Edwards et al. 2005, 2010; Towle et al. 2008), these studies did not observe gamma

suppression following excitation. The reason for the difference may reflect the passive paradigm used in the present study. In primate models of gamma activity in the auditory and visual cortices, gamma activation is associated with increases in single and multiunit firing (Friedman-Hill et al. 2000; Frien et al. 2000; Brosch et al. 2002; Steinschneider et al. 2008). Therefore, it is reasonable to assume that suppression of gamma activity is likely associated with, at the very least, a relative net absence of increased neuronal firing if not outright inhibition. Thus, present findings suggest that after PLST is activated for several hundred milliseconds, it is followed by a relative suppression of activity in the passive awake state.

As a rule, maximal gamma augmentation overlaps the latter half of the N α and leading half of the P β components of the AEP (Crone et al. 2001; Edwards et al. 2005). Both these studies, however, concluded that the dominant generators of the AEP recorded from PLST were from tangential sources located on HG and the superior temporal plane, based upon inversion of AEP components below the Sylvian fissure. Smaller contributions from local radially oriented sources on the lateral surface were also suggested as generators of the AEP recorded from PLST. Similarly, we also observed large amplitude AEPs concurrently recorded with gamma activity from PLST. While a detailed discussion of the AEPs recorded from PLST is beyond the scope of this paper, polarity of the peaks in the AEP recorded below the Sylvian fissure and overlying PLST (e.g., ~100 to 150 ms for the large amplitude N α component, see Figs 6 and 10) are not inverted in polarity from those recorded from the dorsolateral scalp and therefore are not consistent with a dipole source located on the supratemporal plane (Scherg et al. 1989; Liégeois-Chauvel et al. 1994; Yvert et al. 2005). Instead, these responses likely are dominated by locally produced field potentials and are consistent with lateral generators described as the "T" complex (e.g., Celesia 1976; Näätänen and Picton 1987; Knight et al. 1988; Cacace et al. 1990).

The utility of a unified gamma measure (i.e., average ERBP change from baseline across the ECoG frequencies of 75–175 Hz) for examining auditory cortical processing is emphasized by the specificity of the response occurring within the first 200 ms following stimulus onset. Maximal activity was restricted to a confined region of several centimeters centered within PLST. At present, it is unclear what the homology is between the organization of core, belt, and parabelt regions of auditory cortex defined in the old world monkey and that in humans (Kaas and Hackett 2000; Rauschecker and Tian 2000; Hackett et al. 2001; Hackett 2003, 2007). The presence of 1) very short-latency activation of PLST by electrical stimulation of posteromedial HG, 2) short-latency large amplitude AEPs (Howard et al. 2000; Brugge et al. 2003; Reale et al. 2007; Sinai et al. 2009; Chang et al. 2010), 3) short-latency gamma activity (Crone et al. 2001; Edwards et al. 2005; Towle et al. 2008; Sinai et al. 2009) elicited by both speech and nonspeech sounds in this region, and 4) syllable-elicited activity that appears to reflect stimulus acoustics (e.g., stimulus VOT represented in the AEP) instead of higher order speech-related functions strongly suggest that this auditory area is located "early" in the putative human auditory cortex hierarchical processing stream. Although PLST is positioned several centimeters away from core cortex of medial HG, some of these properties seem characteristic of belt cortex, at least as defined physiologically in the monkey model. In a previously reported fMRI study, investigators

observed patterns of auditory evoked responses to speech and nonspeech stimuli within PLST that led them to conclude that this was an “intermediate area” that likely functioned as a belt region of auditory cortex (Uppenkamp et al. 2006). Additional research is clearly required to clarify this issue.

Despite its proposed “intermediate” role, PLST appears to be crucial for speech perception. There is a high specificity between electrode sites on PLST that exhibit large amplitude gamma responses elicited by syllables or words and disruption of speech perception by electrical stimulation (Towle et al. 2008; Sinai et al. 2009). This region also has reciprocal connections, identified through electrical stimulation, with anterior language regions (e.g., Broca’s area, Matsumoto et al. 2004). Stimulation of sites on PLST that had functional connections with anterior language areas could also produce speech arrest. Finally, the potential importance of PLST for speech perception is exemplified by its response sensitivity to syllables varying in consonant POA and VOT.

The presence of high-frequency gamma bands should not be interpreted as a demonstration that all the bands are generated by similar physiological processes. Multiple generators have been proposed for “induced” gamma activity not directly incorporated into the high-frequency phase-locked components of the AEP. These include local circuit interactions among and between pyramidal cells and inhibitory interneurons, all mediated by both synaptic activity and electrical coupling (e.g., Metherate and Cruikshank 1999; LeBeau et al. 2003; Cunningham et al. 2004; Traub et al. 2005; Fries et al. 2007). Therefore, our measure of gamma responses should be viewed not as an index of one specific neuronal process, but as a reliable and sensitive indicator of local cortical activity.

Finally, subjects in this and all studies that involve placement of intracranial electrodes have brain dysfunction. Some patients with temporal lobe epilepsy have language dysfunction, and there is evidence of reorganization of the neural networks subserving speech perception in this setting (e.g., Boatman and Miglioretti 2005; Thivard et al. 2005; Hamberger 2007; Powell et al. 2007). Thus, caution must be exercised when extrapolating findings to subjects without neurologic dysfunction. Generally, intracranial studies examining auditory and language functions incorporate data from a limited number of subjects and from nonoptimal electrode placements determined solely on clinical grounds. Atypical results from even a single subject can thus skew statistical analyses and data interpretation. Given this scenario, it is advantageous for each new study to both replicate and extend previous observations. Replication provides reassurance that experimental findings are reliable indices of normal cortical function. The concordance between this and previous studies on the fundamental observations of timing and spatial distribution of gamma activity overlying the lateral STG indicates their reliability despite variations in subjects, research paradigms, and methods of ECoG analysis and supports the validity of the additional results that make up the remainder of this report.

Outward Expansion of Activation along PLST

Timing of gamma activity is not uniform along the PLST. Gamma augmentation occurs first in a circumscribed region of PLST and is followed by augmentation at both adjacent sites on PLST and in both anterior and posterior directions along the STG. While sequential activation of gamma activation has been

shown during active word processing from PLST to regions outside the STG (Canolty et al. 2007; Edwards et al. 2010), these earlier and less pronounced dynamics have not been previously reported. It is tempting to propose that this progression of activation over time represents the initial stages of an outward expansion of intracortical sound processing along the temporal lobe into adjacent auditory cortical fields that can be observed even in the passive awake state. Similar long-range anterior and posterior expansions have been observed in functional neuroimaging studies examining cortical activation as the intelligibility of speech improved (e.g., Scott et al. 2006; Spitsyna et al. 2006). The currently described outward expansion of activation is plausible based upon anatomically defined corticocortical connections between adjacent auditory fields in the primate (e.g., Kaas and Hackett 2000, Jones 2003). However, this activation pattern could also represent, at least in part, parallel inputs from subcortical areas or other noncontiguous cortical areas (e.g., auditory core). In support of the latter suggestion, data combining fMRI, Granger causality mapping, and anatomical pathway analysis using diffusion tensor probabilistic mapping have shown 2 pathways emanating from primary auditory cortex to both more anterior and posterior regions on the STG (Upadhyay et al. 2008). Further research incorporating multiple methodologies, including tract tracing using electrical stimulation (e.g., Brugge et al. 2003; Matsumoto et al. 2004; Greenlee et al. 2007) will likely be required to resolve this issue.

Representation of Consonant POA

Gamma responses are modulated by consonant POA. There is a statistically significant relationship in the patterns of gamma augmentation 100–150 ms following stimulus onset between responses elicited by voiced CV syllables (/ba/, /da/, and /ga/) and their unvoiced counterparts (/pa/, /ta/, and /ka/), as a function of POA. Thus, one can predict, for instance, that if the response at a given electrode site to the voiced bilabial stop CV syllable /ba/ is larger than /ga/ or /da/, then a similar relationship will hold when examining the relative amplitudes of responses to the unvoiced stops (i.e., /pa/ will elicit the largest response). Furthermore, correlations exist between the relative differences in amplitudes for /ba/ versus /da/ and their unvoiced counterparts /pa/ and /ta/ in the same time frame. An additional correlation between the relative differences in amplitudes for /ga/ versus /da/ and their unvoiced counterparts /ka/ and /ta/ was observed. These findings complement the first intracranial electrophysiological demonstration of an organization of human auditory cortex that is modulated by the phonetic parameter of stop consonant POA (Chang et al. 2010). In that study, the authors identified distinct patterns of AEPs recorded across the posterior lateral surface of the STG in passively listening subjects that respected the phonetic differences between the voiced CV syllables /ba/, /ga/, and /da/. Here, we show complementary ERBP data demonstrating correlations respecting POA despite variation in the voicing parameter of the stop consonants.

Modulation of gamma activity in the 100- to 150-ms time period is consistent with both the work of Chang et al. (2010) and previous noninvasive studies examining the representation of POA in auditory cortex. Chang et al. (2010) found that the 110- to 150-ms time period in the AEP allowed the most accurate discrimination among the stop consonants. Numerous

studies support the importance of early neural activity within several hundred milliseconds bilaterally localized to HG, the planum temporale and PLST for extracting phonetic information from acoustic attributes (for reviews, see Obleser and Eisner 2008; Poeppel et al. 2008). Different dipole source localizations for the “center of gravities” of activity for the magnetic evoked responses P50m and N100m are present for syllables varying in their consonant POA (Obleser et al. 2006; Tavabi et al. 2007). While magnetic responses are relatively insensitive to gyral crest patterns of activity and would thus be relatively unreliable in detecting changes in neural activity on PLST, their timing indicates that information related to consonant POA should be available in the 100- to 150-ms time interval.

The likely relationship between gamma augmentation and increased cellular firing suggests that stop consonant POA is at least partly represented in the PLST by a rate code of neural activity. Temporally discrete rate codes for neural activity evoked by syllable onsets are capable of discriminating responses elicited by stop consonants varying in their POA in A1 of awake monkeys and anesthetized rats (Steinschneider et al. 1995; Engineer et al. 2008; Steinschneider and Fishman 2010). These discriminations are based on the differential spectral content of the syllables at onset and the underlying tonotopic organization. The basis of this relationship is parsimonious with the hypothesis that the short-term spectra within the first 20 ms of consonant onset (release) is a major determinant underlying the perception of stop consonant POA (Stevens and Blumstein 1978; Blumstein and Stevens 1979, 1980; Chang and Blumstein 1981). Overall, this scheme is consistent with findings that vocalizations are represented, at least in A1, by the spectrotemporal discharge pattern of spatially distributed neuron populations determined by the field’s tonotopic organization (e.g., Wang et al. 1995).

While it is attractive to suggest that the selectivity of gamma augmentation in PLST to syllables varying in their consonant POA is based on some underlying spectral organization, other explanations are possible. These include a phonetically based organization (e.g., Chang et al. 2010) or one determined by complex acoustic features such as sensitivity to FM that tracks formant transitions. However, auditory cortex on HG is tonotopically organized (Howard et al. 1996; Formisano et al. 2003; Bitterman et al. 2008), and it would be reasonable to suggest that information transfer from human core auditory cortex to surrounding auditory regions is modulated by this organization. This idea is supported by the tonotopic organizations of belt regions observed in monkeys (Morel et al. 1993; Kosaki et al. 1997; Rauschecker and Tian 2000, 2004). Further support is provided by fMRI evidence that this region differentiates syllables varying along consonant POA based on the spectral content occurring at consonant onset (Obleser et al. 2007). Clearly, detailed investigations of underlying organizational schemes based on acoustic or higher order attributes and their possible relationships with speech-specific activity in the PLST is required (e.g., see Liebenthal et al. 2005).

Finally, it is of note that differentiation of gamma activity respecting consonant POA was obtained using a simple ranking metric based on activity from sites where at least 4 of the 6 syllables elicited responses significantly greater than baseline (see also Steinschneider and Fishman 2010 for a similar approach in monkey A1). It is likely that incorporating additional sites in the analysis or examining response patterns generated across the entire electrode grids may have enhanced

response differentiation. The latter paradigm has been successfully implemented by Chang et al. (2010) and in an fMRI study that used multivariate pattern analysis to classify, and later predict, responses specific for vowel or speaker identity (Formisano et al. 2008). Interestingly, the authors determined that spectral characteristics of the first 2 formants were crucial for vowel differentiation, while the fundamental frequency of the speech sounds was important for classifying speakers. These findings further emphasize the need to investigate PLST for underlying organizations based on spectral, AM, and FM sensitivities. Furthermore, response differences occurred despite the use of highly stylized, and “minimalistic” synthetic syllables whose spectral content and VOT were highly overlapping. All syllables shared the same first and forth formants, /ga/ and /ka/ shared the same second formant transitions as /da/ and /ta/ and the same third formant transitions as /ba/ and /pa, and VOTs were kept constant at either 5 or 40 ms instead of modulated according to consonant POA. It is likely that more realistic syllables with additional differences in their spectral and temporal characteristics would have enhanced response differentiation among the sounds.

Representation of VOT

The physiological representation of the VOT speech parameter has been widely studied in A1 of animal models, intracranial recordings in HG, and human noninvasive recordings (e.g., Eggermont 1995; Schreiner 1998; Simos et al. 1998; Liégeois-Chauvel et al. 1999; Steinschneider et al. 1999, 2003, 2005; Sharma and Dorman 2000; Eggermont and Ponton 2002; Trébuchon-Da Fonseca et al. 2005; Engineer et al. 2008). Animal and human intracranial studies in HG have clearly shown that when the VOT is sufficiently different from 0 ms, there are 2 temporally discrete responses time-locked to consonant release and voicing onset. More detailed analysis of AEPs recorded from the posterior-medial half of HG demonstrates that there is not a single temporal pattern for short and long VOTs (Steinschneider et al. 1999, 2005). Instead, there is a varying pattern with more posterior-medial sites requiring a longer VOT to exhibit a “double-on” response while more anterior-lateral sites required a shorter VOT to exhibit the same temporal response. It was suggested that this variation was based on the interaction between the spectrotemporal features of the syllables and the proposed tonotopic gradient in HG and that the averaged pattern across auditory core would better approximate the perceptual boundary (Steinschneider et al. 2005).

Electrodes on PLST display “single-on” and “double-on” patterns similar to those seen in HG, indicating that VOT can be represented in this auditory area by a temporal code. It remains for further investigation whether this temporal code can predict perceptual boundaries or whether it simply represents the temporal envelope of the sounds. In a similar vein, it also remains for further investigation whether this pattern is more prominent overlying language-dominant cortex (e.g., Liégeois-Chauvel et al. 1999; Trébuchon-Da Fonseca et al. 2005). Even though recording sites illustrated in Figure 10 were in the right hemisphere for both subjects, subject R139 was right hemisphere dominant and subject R149 had bilateral language representation (Table 1).

While temporal representations are available at the level of PLST to differentiate voiced from unvoiced stop consonants,

evidence consistent with a “rate code” was also observed. When, for instance, the gamma response in the intervals from 100 to 150 and 150 to 200 ms was larger for /ba/ than /pa/, a similar relationship was seen for the response differences between /da/ and /ta/ and /ga/ and /ka/. Thus, there appears to be parallel representation, manifested in gamma response amplitude, of both the POA and VOT speech parameters. Once again, while it is tempting to speculate that these representations are based on phonetic categories, explanations based on more fundamental sound attributes are possible. For the VOT parameter, this could be based on differential sensitivity of the recording sites to spectral content. The unvoiced stop consonants contain a greater proportion of higher frequencies embedded in the aspiration noise, which are diminished in the syllables with a VOT of 5 ms. Similar observations were seen in monkey A1 at higher best frequency sites (Steinschneider et al. 2003).

Additional Considerations

Results were acquired under passive listening conditions. Attention to sound modulates scalp-recorded AEPs at early stages of auditory cortical processing, as evidenced by response enhancements beginning at the level of middle latency components and extending into the later N1 and P2 waves (Woldorff and Hillyard 1991; Woldorff et al. 1993; Woods et al. 1994; Neelon et al. 2006; Sabri et al. 2006). Gamma augmentation is selectively enhanced in the auditory and somatosensory cortices when subjects perform simple discriminative tasks (Ray, Niebur, et al. 2008). Thus, future studies will be required to determine the degree to which speech-elicited ERBP changes reflecting stimulus POA and VOT are modulated when subjects perform language-related tasks.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>.

Funding

National Institutes of Health: National Institute on Deafness and other Communication Disorders (R-01-DC04290 to M.H., R-01-DC00657 to M.S.); General Clinical Research Centers Program (MO1-RR-59 to M.H.); Hoover Fund; Carver Trust.

Notes

The authors thank Dr Richard A. Reale and Haiming Chen for assistance in data acquisition and analysis, Carol Dizack for graphic art work, and Dr Daniel Tranel, Dr Robert Jones, and Kenneth Manzel for providing neuropsychological testing data. *Conflict of Interest*: None declared.

References

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Stat Methodol.* 57:289–300.

Besle J, Fischer C, Bidet-Caulet A, Lecaigard F, Bertrand O, Giard M-H. 2008. Visual activation and audiovisual interactions in the auditory cortex during speech perception: intracranial recordings in humans. *J Neurosci.* 28:14301–14310.

Bidet-Caulet A, Fischer C, Bauchet F, Aguera P-E, Bertrand O. 2008. Neural substrate of concurrent sound perception: direct electrophysiological recordings from human auditory cortex. *Front Hum Neurosci.* 1:5.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex.* 10:512–528.

Bitterman Y, Mukamel R, Malach R, Fried I, Nelken I. 2008. Ultra-fine frequency tuning revealed in single neurons of human auditory cortex. *Nature.* 451:197–201.

Blumstein SE, Stevens KN. 1979. Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *J Acoust Soc Am.* 66:1001–1017.

Blumstein SE, Stevens KN. 1980. Perceptual invariance and onset spectra for stop consonant vowel environments. *J Acoust Soc Am.* 67:648–662.

Boatman DF, Miglioretti DL. 2005. Cortical sites critical for speech discrimination in normal and impaired listeners. *J Neurosci.* 25:5475–5480.

Boatman-Reich D, Franaszczuk PJ, Korzeniewska A, Caffo B, Ritzl EK, Colwell S, Crone NE. 2010. Quantifying auditory event-related responses in multichannel human intracranial recordings. *Front Comp Neurosci.* 4:4.

Boemio A, Fromm S, Braun A, Poeppel D. 2005. Hierarchical and asymmetrical temporal sensitivity in human auditory cortices. *Nat Neurosci.* 8:389–395.

Brosch M, Budge E, Scheich H. 2002. Stimulus-related gamma oscillations in primate auditory cortex. *J Neurophysiol.* 87:2715–2725.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA III. 2009. Coding of repetitive transients by auditory cortex on Heschl’s gyrus. *J Neurophysiol.* 102:2358–2374.

Brugge JF, Volkov IO, Garell C, Reale RA, Howard MA, III. 2003. Functional connections between auditory cortex on Heschl’s gyrus and on the lateral superior temporal gyrus in humans. *J Neurophysiol.* 90:3750–3763.

Brugge JF, Volkov IO, Oya H, Kawasaki H, Reale RA, Fenoy A, Steinschneider M, Howard MA III. 2008. Functional localization of auditory cortical fields of human: click-train stimulation. *Hear Res.* 238:12–24.

Cacace AT, Satya-Murti S, Wolpaw JR. 1990. Human middle-latency auditory evoked potentials: vertex and temporal components. *Electroenceph Clin Neurophysiol.* 77:6–18.

Canolty RT, Soltani M, Dalal SS, Edwards E, Dronkers NF, Nagarajan SS, Kirsch HE, Barbaro NM, Knight RT. 2007. Spatiotemporal dynamics of word processing in the human brain. *Front Neurosci.* 1:185–196.

Celesia GG. 1976. Organization of auditory cortical areas in man. *Brain.* 99:403–414.

Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. 2010. Categorical speech representation in human superior temporal gyrus. *Nat Neurosci.* 13:1428–1433.

Chang S, Blumstein SE. 1981. The role of onsets in perception of stop consonant place of articulation: effects of spectral and temporal discontinuity. *J Acoust Soc Am.* 70:39–44.

Crone NE, Boatman D, Gordon B, Hao L. 2001. Induced electrocorticographic gamma activity during auditory perception. *Clin Neurophysiol.* 112:565–582.

Crone NE, Sinai A, Korzeniewska A. 2006. High-frequency gamma oscillations and human brain mapping with electrocorticography. In: Neuper C, Klimesch W, editors. *Progress in brain research.* Vol. 159. Amsterdam (Netherlands): Elsevier. p. 275–295.

Cunningham MO, Whittington MA, Bibbig A, Roopun A, LeBeau FEN, Vogt A, Monyer H, Buhl EH, Traub RD. 2004. A role for fast rhythmic bursting neurons in cortical gamma oscillations in vitro. *Proc Natl Acad Sci U S A.* 101:7152–7157.

Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. 2005. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. *J Neurophysiol.* 94:4269–4280.

Edwards E, Soltani M, Kim W, Dalal SS, Nagarajan SS, Berger MS, Knight RT. 2009. Comparison of time-frequency responses and the event-related potential to auditory speech stimuli in human cortex. *J Neurophysiol.* 102:377–386.

Edwards E, Nagarajan SS, Dalal SS, Canolty RT, Kirsch HE, Barbaro NM, Knight RT. 2010. Spatiotemporal imaging of cortical activation during verb generation and picture naming. *Neuroimage.* 50:291–301.

- Eggermont JJ. 1995. Neural correlates of gap detection and auditory fusion in cat auditory cortex. *NeuroReport*. 6:1645-1648.
- Eggermont JJ, Ponton CW. 2002. The neurophysiology of auditory perception: from single units to evoked potentials. *Audiol Neurootol*. 7:71-99.
- Engineer CT, Perez CA, Chen YH, Carraway RS, Reed AC, Shetake JA, Jakkamsetti V, Chang KQ, Kilgard MP. 2008. Cortical activity patterns predict speech discrimination ability. *Nat Neurosci*. 11:603-608.
- Formisano E, De Martino F, Bonte M, Goebel R. 2008. Who" is saying "what"? Brain-based decoding of human voice and speech. *Science*. 322:970-973.
- Formisano E, Kim D-S, Di Salle F, van de Moortele P-F, Ugurbil K, Goebel R. 2003. Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron*. 40:859-869.
- Friedman-Hill S, Maldonado PE, Gray CM. 2000. Dynamics of striate cortical activity in the alert macaque: I. Incidence and stimulus dependence of gamma-band neuronal oscillations. *Cereb Cortex*. 10:1105-1116.
- Frien A, Eckhorn R, Bauer R, Woelbern T, Gabriel A. 2000. Fast oscillations display sharper orientation tuning than slower components of the same recordings in striate cortex of the awake monkey. *Eur J Neurosci*. 12:1453-1465.
- Fries P, Nikolić D, Singer W. 2007. The gamma cycle. *Trends Neurosci*. 30:309-316.
- Fullerton BC, Pandya DN. 2007. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *J Comp Neurol*. 504:470-498.
- Galaburda A, Sanides F. 1980. Cytoarchitectonic organization of the human auditory cortex. *J Comp Neurol*. 190:597-610.
- Greenlee JD, Oya H, Kawasaki H, Volkov IO, Severson MA, Howard MA, III, Brugge JF. 2007. Functional connections within the human inferior frontal gyrus. *J Comp Neurol*. 503:550-559.
- Hackett TA. 2003. The comparative anatomy of the primate auditory cortex. In: Ghazanfar AA, editor. *Primate audition: ethology and neurobiology*. Boca Raton (FL): CRC Press. p. 199-219.
- Hackett TA. 2007. Organization and correspondence of the auditory cortex of humans and nonhuman primates. In: Kaas JH, editor. *Evolution of the nervous system*. Oxford: Elsevier. p. 109-119.
- Hackett TA. 2008. Anatomical organization of the auditory cortex. *J Am Acad Audiol*. 19:774-779.
- Hackett TA, Preuss TM, Kaas JH. 2001. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol*. 441:197-222.
- Hall DA, Johnsrude IS, Haggard MP, Palmer AR, Akeroyd MA, Summerfield AQ. 2002. Spectral and temporal processing in human auditory cortex. *Cereb Cortex*. 12:140-149.
- Hamberger MJ. 2007. Cortical language mapping in epilepsy: a critical review. *Neuropsychol Rev*. 17:477-489.
- Hart HC, Palmer AR, Hall DA. 2004. Different areas of human non-primary auditory cortex are activated by sounds with spatial and nonspatial properties. *Hum Brain Map*. 21:178-190.
- Herrmann CS, Munk MHJ, Engel AK. 2004. Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn Sci*. 8:347-355.
- Howard MA, III, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA. 1996. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Res*. 724:260-264.
- Howard MA, III, Volkov IO, Mirsky R, Garell PC, Noh MD, Granner M, Damasio H, Steinschneider M, Reale RA, Hind JE, et al. 2000. Auditory cortex on the human posterior superior temporal gyrus. *J Comp Neurol*. 416:79-92.
- Jensen O, Kaiser J, Lachaux J-P. 2007. Human gamma-frequency oscillations associated with attention and memory. *Trend Neurosci*. 30:317-324.
- Jones EG. 2003. Chemically defined parallel pathways in the monkey auditory system. *Ann NY Acad Sci*. 999:218-233.
- Kaas JH, Hackett TA. 2000. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A*. 97:11793-11799.
- Kaas JH, Hackett TA. 2005. Subdivisions and connections of auditory cortex in primates: a working model. In: König R, Heil P, Budge E, Scheich H, editors. *Auditory cortex: a synthesis of human and animal research*. Mahwah (NJ): Lawrence Erlbaum Assoc. p. 7-25.
- Klatt DH, Klatt LC. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am*. 87:820-857.
- Knight RT, Scabini D, Woods DL, Clayworth C. 1988. The effects of lesions of superior temporal gyrus and inferior parietal lobe on temporal and vertex components of the human AEP. *Electroenceph Clin Neurophysiol*. 70:499-509.
- Kosaki H, Hasikawa T, He J, Jones EG. 1997. Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. *J Comp Neurol*. 386:304-316.
- LeBeau FEN, Traub RD, Monyer H, Whittington MA, Buhl E. 2003. The role of electrical signaling via gap junctions in the generation of fast network oscillations. *Brain Res Bull*. 2:3-13.
- Liebhenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. 2005. Neural substrates of phonemic perception. *Cereb Cortex*. 15:1621-1631.
- Liégeois-Chauvel C, de Graaf JB, Laguitton V, Chauvel P. 1999. Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex*. 9:484-496.
- Liégeois-Chauvel C, Musolino A, Badiet JM, Marquis P, Chauvel P. 1994. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroenceph Clin Neurophysiol*. 92:204-214.
- Matsumoto R, Nair DR, LaPresto E, Najm I, Bingaman W, Shibasaki H, Lüders HO. 2004. Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain*. 126:2316-2330.
- Metherate R, Cruikshank SJ. 1999. Thalamocortical inputs trigger a propagating envelope of gamma-band activity in auditory cortex in vitro. *Exp Brain Res*. 126:160-174.
- Morel A, Garraghty PE, Kaas JH. 1993. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol*. 335:437-459.
- Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. 2005. Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. *Science*. 309:951-954.
- Näätänen R, Picton T. 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiol*. 24:375-425.
- Neelon MF, Williams J, Garell PC. 2006. The effects of attentional load on auditory ERPs recorded from human cortex. *Brain Res*. 1118:94-105.
- Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RAW. 2005. Hemodynamic signals correlate tightly with synchronized gamma oscillations. *Science*. 309:948-951.
- Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA III, Brugge JF. 2009. Temporal envelope of time-compressed speech represented in the human auditory cortex. *J Neurosci*. 29:15564-15574.
- Obleser J, Eisner F. 2008. Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci*. 13:14-19.
- Obleser J, Scott SK, Eulitz C. 2006. Now you hear it, now you don't: transient traces of consonants and their nonspeech analogues in the human brain. *Cereb Cortex*. 16:1069-1076.
- Obleser J, Zimmermann J, Van Meter J, Rauschecker JP. 2007. Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb Cortex*. 17:2251-2257.
- Oya H, Kawasaki H, Howard MA III, Adolphs R. 2002. Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *J Neurosci*. 22:9502-9512.
- Poeppl D, Idsardi WJ, van Wassenhove V. 2008. Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc B Biol Sci*. 363:1071-1086.
- Powell HWR, Parker GJM, Alexander DC, Symms MR, Boulby PA, Wheeler-Kingshott CAM, Barker GJ, Koepp MJ, Duncan JS. 2007. Abnormalities of language networks in temporal lobe epilepsy. *Neuroimage*. 36:209-221.

- Privman E, Nir Y, Kramer U, Kipervasser S, Andelman F, Neufeld MY, Mukamel R, Yeshurun Y, Fried I, Malach R. 2007. Enhanced category tuning revealed by intracranial electroencephalograms in high-order human visual areas. *J Neurosci*. 27:6234-6242.
- Rauschecker JP, Tian B. 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc Natl Acad Sci U S A*. 97:11800-11806.
- Rauschecker JP, Tian B. 2004. Processing of band-pass noise in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol*. 91:2578-2589.
- Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS. 2008. Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci*. 28:11526-11536.
- Ray S, Niebur E, Hsiao SS, Sinai A, Crone NE. 2008. High-frequency gamma activity (80-150 Hz) is increased in human cortex during selective attention. *Clin Neurophysiol*. 119:116-133.
- Reale RA, Calvert GA, Thesen T, Jenison RL, Kawasaki H, Oya H, Howard MA, III, Brugge JF. 2007. Auditory-visual processing represented in the human superior temporal gyrus. *Neurosci*. 145:162-184.
- Recanzone GH. 2008. Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey. *J Neurosci*. 28:13184-13193.
- Rivier F, Clarke S. 1997. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. *Neuroimage*. 6:288-304.
- Sabri M, Liebenthal E, Waldron EJ, Medler DA, Binder JR. 2006. Attentional modulation in the detection of irrelevant deviance: a simultaneous ERP/fMRI study. *J Cogn Neurosci*. 18:698-700.
- Scherg M, Vajsar J, Picton TW. 1989. A source analysis of the late human auditory evoked potentials. *J Cogn Neurosci*. 1:336-355.
- Schreiner CE. 1998. Spatial distribution of responses to simple and complex sounds in the primary auditory cortex. *Audiol Neurootol*. 3:104-122.
- Scott SK, Johnsrude IS. 2003. The neuroanatomical and functional organization of speech perception. *Trends Neurosci*. 26:100-107.
- Scott SK, Rosen S, Lang H, Wise RJS. 2006. Neural correlates of intelligibility in speech investigated with noise vocoded speech—a positron emission tomography study. *J Acoust Soc Am*. 120:1075-1083.
- Sharma A, Dorman MF. 2000. Neurophysiological correlates of cross-language phonetic perception. *J Acoust Soc Am*. 107:2697-2703.
- Simos PG, Diehl RL, Breier JI, Molis MR, Zouridakis G, Papanicolaou AC. 1998. MEG correlates of categorical perception of a voice onset time continuum in humans. *Cogn Brain Res*. 7:215-219.
- Sinai A, Crone NE, Wied HM, Franaszczuk PJ, Miglioretti D, Boatman-Reich D. 2009. Intracranial mapping of auditory perception: event-related responses and electrocortical stimulation. *Clin Neurophysiol*. 120:140-149.
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS. 2006. Converging language streams in the human temporal lobe. *J Neurosci*. 26:7328-7336.
- Steinschneider M, Fishman YI. 2010. Enhanced physiologic discriminability of stop consonants with prolonged formant transitions in awake monkeys based on the tonotopic organization of primary auditory cortex. *Hear Res*. doi:10.1016/j.heares.2010.04.008.
- Steinschneider M, Fishman YI, Arezzo JC. 2003. Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *J Acoust Soc Am*. 114:307-321.
- Steinschneider M, Fishman YI, Arezzo JC. 2008. Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. *Cereb Cortex*. 18:610-625.
- Steinschneider M, Reser D, Schroeder CE, Arezzo JC. 1995. Tonotopic organization of responses reflecting stop consonant place of articulation in primary auditory cortex (A1) of the monkey. *Brain Res*. 674:147-152.
- Steinschneider M, Volkov IO, Fishman YI, Oya H, Arezzo JC, Howard MA, III. 2005. Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cereb Cortex*. 15:170-186.
- Steinschneider M, Volkov IO, Noh MD, Garell PC, Howard MA III. 1999. Temporal encoding of the voice onset time phonetic parameter by field potentials recorded directly from human auditory cortex. *J Neurophysiol*. 82:2346-2357.
- Stevens KN, Blumstein SE. 1978. Invariant cues for place of articulation in stop consonants. *J Acoust Soc Am*. 64:1358-1368.
- Sweet RA, Dorph-Petersen K-A, Lewis DA. 2005. Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *J Comp Neurol*. 491:270-289.
- Tavabi K, Obleser J, Dobel C, Pantev C. 2007. Auditory evoked fields differentially encode speech features: an MEG investigation of the P50m and N100m time courses during syllable processing. *Eur J Neurosci*. 25:3155-3162.
- Thivard L, Hombrouck J, Tézenas du Montcel S, Delmaire C, Cohen L, Samson S, Dupont S, Chiras J, Baulac M, Lehericy S. 2005. Productive and perceptive language reorganization in temporal lobe epilepsy. *Neuroimage*. 24:841-851.
- Tian B, Rauschecker JP. 2004. Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *J Neurophysiol*. 92:2993-3013.
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. 2001. Functional specialization in rhesus monkey auditory cortex. *Science*. 292:290-293.
- Towle VL, Yoon H-A, Castelle M, Edgar JC, Biassou NM, Frim DM, Spire J-P, Kohrman MH. 2008. ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. *Brain*. 131:2013-2027.
- Traub RD, Bibbig A, LeBeau FEN, Cunningham MO, Whittington MA. 2005. Persistent gamma oscillations in superficial layers of rat auditory neocortex: experiment and model. *J Physiol*. 562:3-8.
- Trébuchon-Da Fonseca A, Giraud K, Badier J-M, Chauvel P, Liégeois-Chauvel C. 2005. Hemispheric lateralization of voice onset time (VOT) comparison between depth and scalp EEG recordings. *Neuroimage*. 27:1-14.
- Upadhyay J, Silver A, Knaus TA, Lindgren KA, Ducros M, Kim D-S, Tager-Flusberg H. 2008. Effective and structural connectivity in the human auditory cortex. *J Neurosci*. 28:3341-3349.
- Uppenkamp S, Johnsrude IS, Norris D, Marslen-Wilson W, Patterson RD. 2006. Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage*. 31:1284-1296.
- Wallace MN, Johnston PW, Palmer AR. 2002. Histochemical identification of cortical areas in the auditory region of the human brain. *Exp Brain Res*. 143:499-508.
- Wang X, Merzenich MM, Beitel R, Schreiner CE. 1995. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J Neurophysiol*. 74:2685-2706.
- Woldorff MG, Hillyard SA. 1991. Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroenceph Clin Neurophysiol*. 79:170-191.
- Woldorff MG, Gallen CC, Hampson SR, Hillyard SA, Pantev C, Sobel D, Bloom FE. 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc Natl Acad Sci U S A*. 90:8722-8726.
- Woods DL, Alho K, Algazi A. 1994. Stages of auditory feature conjunction: an event-related brain potential study. *J Exp Psych: Hum Perc Perf*. 20:81-94.
- Woods TM, Lopez SE, Long JH, Rahman JE, Recanzone GH. 2006. Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. *J Neurophysiol*. 96:3323-3337.
- Yvert B, Fischer C, Bertrand O, Pernier J. 2005. Localization of human supratemporal auditory areas from intracerebral auditory evoked potentials using distributed source models. *Neuroimage*. 28:140-153.